Considerations for adaptive tutoring within serious games: authoring cognitive models and game interfaces

Robert A. Sottilare[1] and Stephen Gilbert[2]

[1] Army Research Laboratory – Human Research & Engineering Directorate – Simulation & Training Technology Center
robert.sottilare@us.army.mil
[2] Iowa State University - Human Computer Interaction, Psychology - Virtual Reality Applications Center
gilbert@iastate.edu

Abstract. In reflecting on a recent emphasis on self-directed learning using game-based simulations and virtual worlds, the authors considered key challenges in transforming serious games and virtual worlds into adaptive training tools. This article reflects specifically on the challenges and potential of cognitive modeling techniques for game-based tutors and game integration with computer-based. Finally, as we integrate tutors and games, a rubric is offered to identify some of the salient characteristics of adaptive game-based, tutoring systems.

Keywords: serious games, virtual worlds, intelligent tutoring, cognitive modeling, authoring

1 Introduction

The emergence of "serious games" has sparked interest in their potential as training and educational systems. "Serious games" are a genre of games used in a professional context (e.g., military operations, medical care or emergency management) and generally focus on collective (team-based) tasks. Providing a personalized computer-based tutoring experience for individuals involved in collective training can be challenging. Assessing team cognition and interaction during team training is even more challenging, but is an important part of collaborative and social learning [17]. A good starting point might be to look at how technology (tools and methods) support individual self-directed learning.

Woolf [21] established a set of grand challenges for tutoring system technology that serious games should aspire to if they are to realize their full potential as educational tools. Among others, she notes two key challenges: personalized training and education; and the assessment of learning. The first includes the tailoring and adaptation of instruction based on the learner's capabilities and other pertinent information (a learner model). The second challenge is one faced by every human tutor: How do I recognize/assess if the student is learning and when to adapt the instruction to meet their needs?

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **JUN 2011**	2. REPORT TYPE	3. DATES COVERED **00-00-2011 to 00-00-2011**

4. TITLE AND SUBTITLE **Considerations for adaptive tutoring within serious games: authoring cognitive models and game interfaces**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Army Research Laboratory,Human Research & Engineering Directorate,Simulation & Training Technology Center,Orlando,FL,32826**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Presented at the 15th International Conference Artificial Intelligence in Education, Auckland, New Zealand, June 28-July 1, 2011

14. ABSTRACT
In reflecting on a recent emphasis on self-directed learning using game-based simulations and virtual worlds, the authors considered key challenges in transforming serious games and virtual worlds into adaptive training tools. This article reflects specifically on the challenges and potential of cognitive modeling techniques for game-based tutors and game integration with computer-based. Finally, as we integrate tutors and games, a rubric is offered to identify some of the salient characteristics of adaptive game-based, tutoring systems.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**10**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

In an effort to more easily tailor training and assess learning, this article considers two key capabilities in the evolution of "serious games" as self-directed training tools: authoring cognitive models and game integration with tutors. Today, commercial "serious games" lack cognitive modeling (e.g., learner or expert models) that would allow them to assess/adapt to the learner's needs and computer-based tutoring system interfaces have generally been handcrafted, one-of-a-kind solutions.

Figure 1 illustrates a construct for adaptive, personalized training using game-based scenario templates. Instead of having learners participate in a standardized scenario, e.g., stabilize a wounded soldier and make a medical evacuation request, the scenario would be based on a flexible template with a variety of parameters that can be adjusted depending on macro (pre-training data) and micro (real-time) factors, including but are not limited to: relevant learner skills; individual differences that influence learning (e.g., personality profiles); micro or near-term performance during the scenario; physiological and behavioral data used to classify learner states such as engagement or confusion.

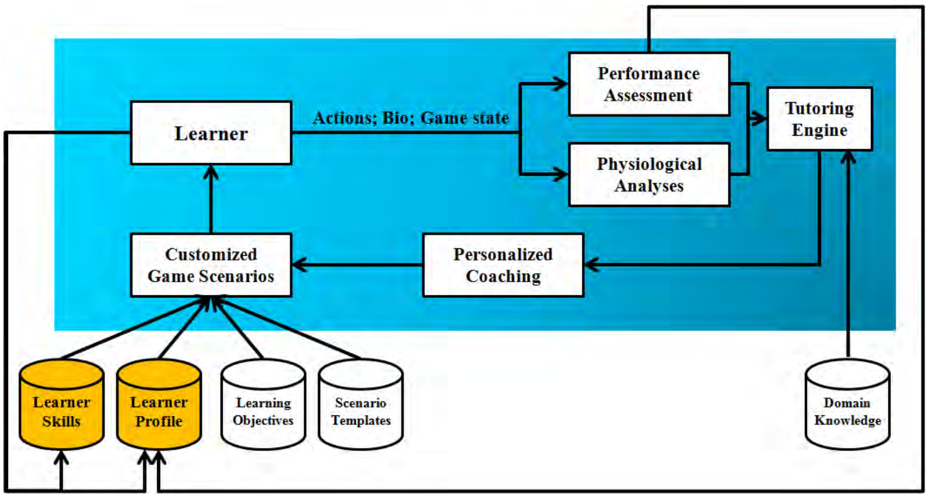

Fig. 1: A construct for personalized adaptive training via game scenario templates, which are customized, based on initial learner skills and a learner profile including performance and cognitive models based on behaviors and physiological measures.

In the figure, assessment is based on both the learner's performance and on physiological data. The performance assessment would be used to make adjustments to the curriculum, tweaking scenarios according to the skills still requiring mastery and ensuring that the learner can be trained in essential tasks in relevant conditions and to established standards. The physiological data would be used to assess the learner's engagement with the system [16] – Is she bored? Is he overwhelmed? This physiological data can also be used, when appropriate, for stress inoculation pedagogies in which the learner is exposed to stressful practice situations so that live stressful situations on the job are handled more smoothly [12]

All of this assessment processing and the generation of personalized appropriate feedback would be handled by the tutoring engine and the domain knowledge,

typically a cognitive model of the learner and an ideal learner, but how will the serious game (where the learner interacts) share state and interaction data with the computer-based tutor? How will the tutor provide feedback and scenario changes to the game? This paper discusses some of the challenges involved in developing just such an engine and an appropriate model. Since serious games are often played in one-to-one or one-to-many contexts without a human tutor, two challenges could be significant in the development of games as training tools: cognitive and affective modeling of the learner based on physiological signals and behaviors from game play; and creating effective authoring tools for game and tutor integration.

2 Cognitive Models and Tutoring Interfaces for Serious Games

Computer-based tutoring systems use expert models (also known as ideal learner models) to compare to learner performance and determine learner progress versus expected progress [3]. Generally, these expert models are painstakingly generated through task analyses. Observers develop a detailed description of behaviors and mental activities, task conditions and standards, and other factors leading to successful performance. An expert model stores information about the instructional content, common questions, common misconceptions, solutions, and expert methodologies. The addition of physiological and behavioral measures to a game-based environment changes the approach to building expert models in two ways.

Even today, computer-based tutors are often one-of-a-kind artifacts that are often handcrafted by experts in a specific training domain (e.g., negotiation training or casualty care). Authoring tools allow non-programmers to create models/content for computer-based tutors and thereby reduce the cost and increase the producibility of intelligent tutoring systems [13]. In addition to producing content, authoring systems also enable the production of models by experts in the domain who may not have programming expertise. Tutoring systems such as AutoTutor [10], CTAT [1], and xPST [8] all have authoring systems that allow model builders with limited programming skills to design tutoring content. Each tool has its pros and cons, and makes compromises on the scale of power vs. ease of use.

However, it is likely that authoring tools for game-based tutors will need to be robust enough to represent many of the objects and actions of the game itself. In order to tutor on location relationships, for example, whether a player is near a particular building, the cognitive model will need to contain a model of the environment that can be queried about that geographic relationship, or perhaps an advanced query function that draws on the game's terrain database directly. This context requires the tutor developer to either understand the inner workings of the game enough to duplicate it or create functions that relate to those inner workings. It is therefore more critical than ever that authoring tools be powerful yet easy to use.

2.1 The Challenge of Including Physiological and Behavioral Measures to Support Cognitive Modeling in Serious Games

The importance of relevant and timely feedback in one-to-one tutoring is well documented [4, 15, 18]. Computer-based systems that provide one-to-one tutoring must also be able to understand the learner's state and adapt scenario challenge-level and flow in support of the training objectives [7]. To this end, it would be useful for game-based tutors to be able to model the learner's unobserved cognitive state including their affect (personality, emotions or mood), their readiness to learn (attention, engagement and motivation) and their comprehension. These unobserved outcomes (affect, readiness to learn and comprehension) may be determined probabilistically by their relationship to observed variables (e.g., physiological measures, behavioral measures, human observations and self-reported information) [15, 17, 18].

A 2010 expert panel of military clinical psychologists [19] described the potentially revolutionary impact on improved stress training and mental health using such personalized adaptive training system, both to each preparation for stressful battle events as well as for treating post-traumatic stress disorder (PTSD) via exposure therapy. One of the therapists, Dr. Scott Johnston of the Naval Medical Center San Diego, described the challenges faced by Marines in Iraq, noting "To help build some resilience with them by training them in appropriate VR scenarios they would likely encounter before leaving on deployment is a very exciting prospect." The growing field of neuroergonomics [14] offers a number of approaches for inferring cognitive states from physiological signals that will need to be incorporated into tutors for strong personalization.

Cognitive modelers, however, who may be accustomed to constructing models that support procedural tutors ("If the learner does step X, then step Y, then step Z, then give feedback F") or tutors in which subgoals are assumed to be equivalent to the completion of actions may not be accustomed to a more state-based approach. In this context, states might be "inter-beat heart rate high" or "experiencing positive stress, gearing up for the challenge." For example, cognitive models will have an added level of complexity over and above production rules such as "while the learner is in physiological state X, the following rules apply; while in state Y or state Z for at least 10 minutes, then the following different rules apply." Cognitive models will expand to contain not only the typical abstracted domain knowledge, but also the learner context around that knowledge (meta-data). Game-based tutors will better be able to adapt to learner needs when they are able to use such complex cognitive models to predict learner cognitive states based on observed variables including behavioral interaction within the game.

2.2 The Challenge of Integrating Serious Games and Tutors

For some cognitive modelers, building tutors for games may also require thinking differently about monitoring the learner's progress through the game. In games and virtual environments, there are frequently many methods of achieving a goal. For example, if a game player wants to move from a "home" location to a "cave" location,

she might walk, jump, run, teleport, etc. If the learning objectives are focused on the methods of locomotion, then including those methods in the cognitive model would be appropriate. However, if the learning objectives focus instead of how to explore caves, and the instructor/cognitive modeler does not care how the cave is reached, then locomotion methods would not be included. In comparison with popular intelligent tutoring systems domains such as mathematics or physics, which typically use a graphic user interface (GUI), game domains present a much broader range of granularity of events and stimuli that can be taken into account during learning. Since model-tracing tutors [2] focus on specifically on matching patterns of events to elements of the model, cognitive modelers will have choices about the level of granularity to focus on.

Game environments also typically add additional complexity based on the additional variables involved in playing. Is the player carrying the right items at the right moment? Does the player have the right combination of health status, strength attributes, and collected items to resolve the next challenge? Secondly, given that many more complicated games allow players to build their in-game talents or skills, an instructional designer might ask the extent to which in-game within-character skills correspond to real-world player skills (transfer). In games such as World of Warcraft, "grinding" behaviors are popular (boring, highly repetitive tasks that accumulate rewards and talent levels). How might a tutor interact with players' interest in such activities?

Annie [20] is an example of a system for embedding computer-based tutors into games. It guides learners through a computer game by using the game's core mechanics. Annie represents initial/goal states of learner knowledge and provides guidance to support learning. Game difficulty and learner skill level are evaluated with the goal of maintaining the learner in the "optimal gameplay corridor" where boredom and confusion are avoided. However, Annie does not specifically classify learner affect (e.g., emotions like boredom or confusion) to adapt the tutor's actions. The Extensible Problem-Specific Tutor (xPST) has also made forays into tutoring on games, offering the opportunity to build tutors around location-based event triggers and actions by multiple entities [9].

The existence of even these two game-oriented tutoring approaches suggests the need for an agreed-upon abstraction layer that could be used to interface between tutoring engines and game engines (see Figure 2). Such an abstraction would enable tutors to be written independently of the game in which they are tutoring. Cognitive modelers could create tutors that refer to players, non-player characters (NPCs), or other objects, and the abstraction layer would translate those objects to the appropriate world within game engines such as VBS2, Ogre, or Unity, which would normally require an author to use their respective game engine scripting languages. Simulation APIs such as DEVS (Discrete Event System Specification) created by Zeigler [22] or BrahmsVE [5] may be useful role models for such an abstraction layer.

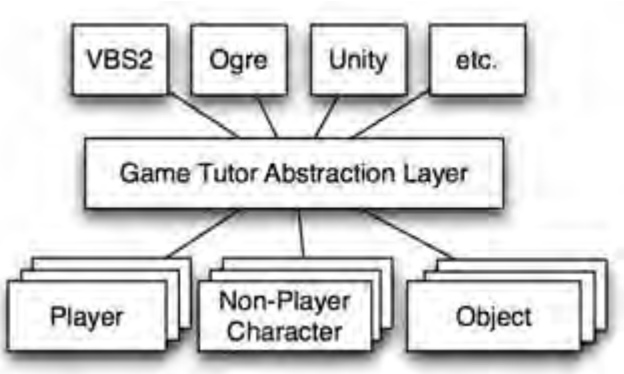

Fig. 2: A game tutor abstraction layer would provide a standard set of protocols that would enable a tutor to work with multiple game engines.

3 Discussion

Given the state of practice of serious games, what research is needed to optimize their training potential? Serious games (and virtual worlds) have some serious deficiencies as training tools, but also significant potential if they can be easily linked to tutoring systems and if they are able to adapt instruction to individual learners. Some of the limitations and current capabilities of serious games as training tools are discussed below.

The adaptability of serious games is constrained by their limited artificial intelligence (AI) and lack of cognitive modeling to support adaptive training. In the area of cognitive modeling, serious games could be greatly improved by the addition of learner modeling standards to improve the portability of learner models to other games, virtual worlds and other training/educational simulations. The authors recommend additional research to improve the real-time understanding of the learner's cognitive state (e.g., engagement level, emotions) through unobtrusive (passive) behavioral and physiological sensing (neuroergonomics). The classification of the learner's cognitive state might also be improved by access to historical, self-reported and observer data (e.g., competency measures, social learning profiles and preference surveys) made accessible through online learning management systems. However, improving cognitive modeling will not improve learning substantially unless instructional strategy (e.g., scenario adaptation, feedback, hints, motivational support) classification models are optimized. The authors also advocate additional research in the development of cognitive models for teams to support adaptive tutoring beyond the one-to-one training focus of today's research. This would make tutors more compatible with play in serious games which is mostly collective (team-focused).

Authoring continues to head the list of many in the tutoring system and distributed learning domains. Sometimes authoring is enabled in the form of standards like the Shareable Content Object Reference Model (SCORM) which promotes the reuse of

courseware. The use of computer-based tutors that can be implemented across serious games and interact with them like Annie [20], promote reuse and have the potential to reduce development and support costs for training.

Authoring tools to simplify the integration of tutors and serious games would also help tutors more easily assess the learner's cognitive state. Figure 3 shows how this might work. The abstraction layer discussed earlier could be composed of both a game interface layer and a tutor interface layer. Since many serious games have distributed simulation standard protocol interfaces (e.g. Distributed Interactive Simulation or High Level Architecture), it might simplify interaction by adopting one of these standards for the tutor interface layer. Data transfer between these layers might be limited to five essential data types: entity state data, game state data, interaction data, non-player character (NPC) feedback and scenario changes.

Entity state data is composed of things like the entity type (e.g., person, vehicle) and associated information about their location, position and behavioral events (e.g., decisions). Interaction data includes physical and social interaction between entities. This behavioral data (entity and interaction) may be influential in optimizing instructional strategies (e.g., feedback, changes in flow and challenge level). Game state data represents the physical training environment.

The tutor uses the entity, interaction and game state data along with sensor and historical performance data to assess, model, predict and influence the selection of instructional strategies (e.g. feedback or change in flow or challenge level).

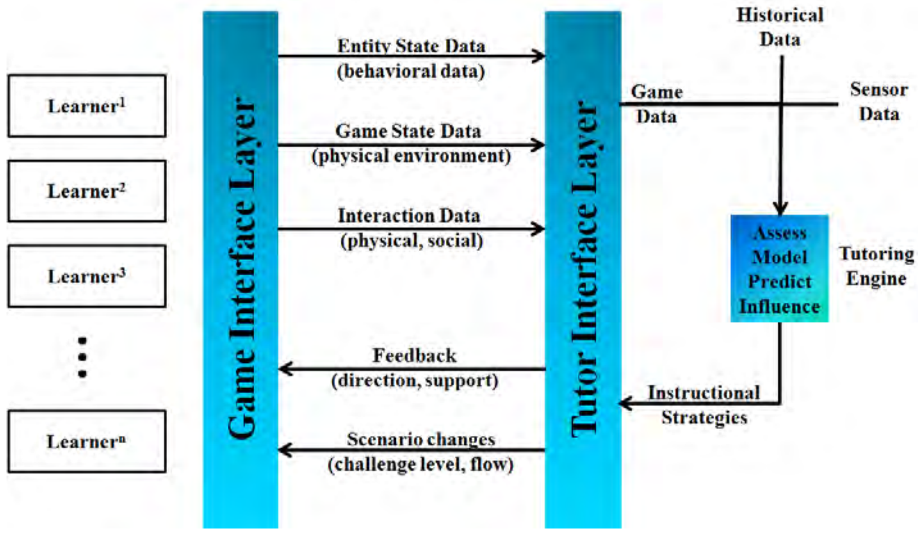

Fig. 3: A game-tutor abstraction layer to extract relevant behavioral data from games for consumption by tutors for instruction strategy decisions (feedback and scenario changes) that are actionable in the game.

4 Conclusions

It is a significant benefit to think that many existing games have the potential to be training systems, but few have sufficient content to support specific training objectives, real-time feedback or measures of performance. Additional research is needed to support improve processes for cognitive model development for both learners (so we can understand and adapt to their state) and expert models (so we can compare learner performance to an “ideal standard”).

In the area of authoring tools, additional research is needed for “intelligent authoring” including the automated development of cognitive (learner and expert) models; automated development/adaptation of scenario content based on learner needs and standard frameworks for authoring and interacting with training content.

As research moves forward and the tutoring system community continues to progress toward an “ideal computer-based tutor”, a measuring stick is needed to assess the effectiveness of adaptive computer-based tutors. The authors offer the four categories noted below as a tutoring yardstick based on the premise that a “gold standard” is to have tutoring systems of equal quality and adaptability as a proficient human tutor in one-to-one training environments. This human equivalence has also been described by Bloom [4] as the two-sigma (2σ) difference that has been demonstrated to exist between a learner's achievements in a classroom vs. a learner's achievement with a one-on-one tutor. Kulik [11] surveyed 97 research studies on tutors and found that most tutoring systems have an average difference (or "effect size") of 0.32σ. Cognitive Tutors have reached an effect size of up to 1σ [6], but these were not game-based tutors. The authors defined four levels of adaptive tutors as a set of goals for tutor-developers and measures of effectiveness for consumers of tutoring products. While primarily developed to assess game-based tutors, there would be no impediment to their use in other tutoring contexts.

Platinum Tutors ($> 2\sigma$ difference from traditional classroom): are able to adapt to the learner better than a human tutor; enable learning better than a human tutor; fully perceive learner behaviors and physiology through remote sensing; support fully mobile training; are consistently accurate (near 100%) in classifying the learner’s cognitive state in near real-time; have an optimized repertoire of instructional strategies; and finally, are automatically integrated with a variety of training platforms (e.g., serious games, commercial/military training simulations).

Gold Tutors (2σ difference from traditional classroom): are able to adapt to the learner equally as well as a human tutor; enable learning equal to the best human tutors; fully perceive learner behaviors through unobtrusive sensing methods; support mobile training within instrumented spaces; are consistently accurate (>80%) in classifying the learner’s cognitive state in near real-time; have an optimized repertoire of instructional strategies; and finally, are easily integrated by laypeople with a variety of training platforms (e.g., serious games, commercial/military training simulations).

Silver Tutors (approximately 1σ-2σ difference): are able to adapt to the learner nearly as well as a human tutor; enable learning equal to good human tutors; fully perceive learner behaviors through unobtrusive sensing methods; are not mobile (static classroom systems); are consistently accurate (>80%) in classifying the learner’s cognitive state; use prescriptive instructional strategies; and finally, are

easily integrated by professionals with a variety of training platforms (e.g., serious games, commercial/military training simulations).

Bronze Tutors (< 1σ difference; many tutors today): adapt to learner only in prescribed ways based on an intrinsic expert model; offer improvements on self-directed learning with no tutor; perceive some learner behaviors primarily related to performance; are static classroom-based systems; can classify the learner's cognitive state only based on performance and do not account for learner state; use prescriptive instructional strategies; and finally, may be integrated by professionals with a limited class of training platforms (e.g., serious games).

References

1. Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. (2006). The Cognitive Tutor Authoring Tools (CTAT): Preliminary evaluation of efficiency gains. In M. Ikeda, K. D. Ashley, & T. W. Chan (Eds.), Proceedings of the 8th International Conference on Intelligent Tutoring Systems (ITS 2006), (pp. 61-70).
2. Anderson, J.R., & Pelletier, R. (1991). A development system for model-tracing tutors. In L. Birnbaum (Ed.), Proceedings of the International Conference of the Learning Sciences (pp. 1-8). Charlottesville, VA: Association for the Advancement of Computing in Education.
3. Beck, J., Stern, M., and Haugsjaa, E. (1996) Applications of AI in Education, ACM Crossroads 3(1), 11-15.
4. Bloom, Benjamin S. (1984) The 2-sigma problem: The search for methods of group instruction as effective as one-to-one tutoring, Educational Researcher 13: 4-16.
5. Clancey, W. J., Sachs, P., Sierhuis, M., & van Hoof, R. (1998). Brahms: Simulating practice for work systems design. International Journal on Human-Computer Studies, 49, 831-865.
6. Corbett, A.T. (2001) Cognitive Computer Tutors: Solving the Two-Sigma Problem. In Proceedings of the 8th International Conference on User Modeling 2001 (UM '01), Mathias Bauer, Piotr J. Gmytrasiewicz, and Julita Vassileva (Eds.). Springer-Verlag, London, UK, 137-147.
7. Csikszentmihalyi, M. (1990). Flow: the psychology of optimal experience. Harper Perennial, New York, NY.
8. Gilbert, S., Blessing, S. B., Kodavali, S. (2009) The Extensible Problem-Specific Tutor (xPST): Evaluation of an API for Tutoring on Existing Interfaces. Proceedings of the 14th International Conference on Artificial Intelligence in Education.
9. Gilbert, S., Devasani, S., Kodavali, S., Blessing, S. (2011). Easy Authoring of Intelligent Tutoring Systems for Synthetic Environments. Proceedings of the Twentieth Conference on Behavior Representation in Modeling and Simulation.
10. Graesser, A.C., Chipman, P., Haynes, B.C., & Olney, A. (2005) AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. IEEE Transactions in Education, 48, 612–618.
11. Kulik, J. A. (1994). Meta-Analytic studies of findings on computer-based instruction. In E. L. Baker and H. F. O'Neil (Eds.), Technology Assessment in Education and Training (pp. 9-33). Hillsdale, NJ: Lawrence Erlbaum Associates.
12. Meichenbaum, D. (1996). Stress inoculation training for coping with stressors. The Clinical Psychologist, 49, 4-7.
13. Murray, T., Blessing, S., & Ainsworth, S. (Eds.). (2003). Authoring Tools for Advanced Technology Educational Software. Dordrecht, The Netherlands: Kluwer Academic Publishers.

14. Parasuraman, R., & Rizzo, M. (Eds.), (2006) Neuroergonomics: The Brain at Work. Oxford, UK: Oxford University Press.
15. Picard, R. (2006). Building an Affective Learning Companion. Keynote address at the 8th International Conference on Intelligent Tutoring Systems, Jhongli, Taiwan. Retrieved from http://www.its2006.org/ITS_keynote/ITS2006_01.pdf
16. Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. Biological Psychology, 40(1-2), 187-195.
17. Sottilare, R. (2010). Toward the Development of an Intelligent Tutoring System for Distributed Team Training through Passive Sensing. In Proceedings of the Intelligent Tutoring Systems (ITS) 2010 Conference, Pittsburgh, Pennsylvania, June 2010.
18. Sottilare, R. and Proctor, M. (2011). Classifying student mood within intelligent tutoring systems (ITS). Accepted for publication in the Journal of Educational Technology.
19. Spira, et al, 2010. Expert panel: future directions of technological advances in prevention, assessment, and treatment for military deployment mental health. Cyberpsychology, Behavior, and Social Networking, 13(1).
20. Thomas, J. and Young, R. M. (2009). Dynamic Guidance in Digital Games: Using an Extensible Plan-Based Representation of Exploratory Games to Model Student Knowledge and Guide Discovery Learning. In Working Notes of the Intelligent Educational Games Workshop at the International Conference on Artificial Intelligence and Education (AIED 09), Brighton, UK, July, 2009.
21. Woolf, B. P. (2010). A Roadmap for Education Technology. National Science Foundation # 0637190.
22. Zeigler, B. (1976). Theory of Modeling and Simulation (first ed.). Wiley Interscience, New York.